

Published before June 1999

Found 7,095 of 102,899

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Sort results
by

[relevance](#)

 [Save results to a Binder](#)

Try an [Advanced Search](#)

Display
results

[expanded form](#)

 [Search Tips](#)

Try this search in [The ACM Guide](#)

[Open results in a new window](#)

Results 1 - 20 of 200

Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale 

1 Projections for efficient document clustering



Hinrich Schütze, Craig Silverstein

July 1997 **ACM SIGIR Forum , Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '97**, Volume 31 Issue SI

Publisher: ACM Press

Full text available:  [pdf\(1.26 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



2 Using Kohonen maps to determine document similarity

Jennifer Farkas

October 1994 **Proceedings of the 1994 conference of the Centre for Advanced Studies on Collaborative research**

Publisher: IBM Press

Full text available:  [pdf\(575.72 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)



In this paper we present some experimental results on the classification of natural language documents using Kohonen's self-organizing-map neural network paradigm. We discuss, in particular, how the classification accuracy can be improved if the standard keyword representation of documents is enhanced by including specific weights, thesaurally-defined relations among keywords, and additional synonyms for keywords. We sketch the main features of a prototype of an automatic document classification ...

3 Special issue on word sense disambiguation: Automatic word sense discrimination

Hinrich Schütze

March 1998

- 4 [TourisT: the application of a description logic based semantic hypermedia system for tourism](#)

 Joe Bullock, Carole Goble

May 1998 **Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space---structure in hypermedia systems: links, objects, time and space---structure in hypermedia systems**

Publisher: ACM Press

Full text available:  pdf(1.57 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

- 5 [HyPursuit: a hierarchical network search engine that exploits content-link hypertext clustering](#)

 Ron Weiss, Bienvenido Vélez, Mark A. Sheldon

March 1996 **Proceedings of the the seventh ACM conference on Hypertext**

Publisher: ACM Press

Full text available:  pdf(2.00 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

- 6 [Using latent semantic indexing for information filtering](#)

 P. W. Foltz

March 1990 **ACM SIGOIS Bulletin , Proceedings of the ACM SIGOIS and IEEE CS TC-OA conference on Office information systems**, Volume 11 Issue 2-3

Publisher: ACM Press

Full text available:  pdf(894.90 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Latent Semantic Indexing (LSI) is an information retrieval method that organizes information into a semantic structure that takes advantage of some of the implicit higher-order associations of words with text objects. The resulting structure reflects the major associative patterns in the data while ignoring some of the smaller variations that may be due to idiosyncrasies in the word usage of individual documents. This permits retrieval based on the "latent" semantic content of t ...

- 7 [A step towards the detection of semantic variants of terms in technical documents](#)

Thierry Hamon, Adeline Nazarenko, Cécile Gros

August 1998 **Proceedings of the 36th annual meeting on Association for Computational Linguistics - Volume 1 , Proceedings of the 17th international conference on Computational linguistics - Volume 1**

Publisher: Association for Computational Linguistics , Association for Computational Linguistics

Full text available:  pdf(606.64 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)


This paper reports the results of a preliminary experiment on the detection of semantic variants of terms in a French technical document. The general goal of our work is to help the structuration of terminologies. Two kinds of semantic variants can be found in traditional terminologies: strict synonymy links and fuzzier relations like see-also. We have designed three rules which exploit general dictionary information to infer synonymy relations between complex candidate terms. The results ...

- 8 [Use of syntactic context to produce term association lists for text retrieval](#)

 Gregory Grefenstette

June 1992 **Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available: [pdf\(714.72 KB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

One aspect of world knowledge essential to information retrieval is knowing when two words are related. Knowing word relatedness allows a system given a user's query terms to retrieve relevant documents not containing those exact terms. Two words can be said to be related if they appear in the same contexts Document co-occurrence gives a measure of word relatedness that has proved to be too rough to be useful. The relatively recent apparition of on-line dictionaries and robust and rapid par ...

9 The document lens

 George G. Robertson, Jock D. Mackinlay

December 1993 **Proceedings of the 6th annual ACM symposium on User interface software and technology**

Publisher: ACM Press

Full text available: [pdf\(689.28 KB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



Keywords: 3D interaction techniques, graphic presentations, interface metaphors, screen layout, user interface design issues

10 Information retrieval & extraction: N-gram cluster identification during empirical knowledge representation generation

Robin Collier

August 1994 **Proceedings of the 15th conference on Computational linguistics - Volume 2**

Publisher: Association for Computational Linguistics

Full text available: [pdf\(467.30 KB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#)



This paper presents an overview of current research concerning knowledge extraction from technical texts. In particular, the use of empirical techniques during the identification and generation of a semantic representation is considered. A key step is the discovery of useful n-grams and correlations between clusters of these n-grams.

Keywords: knowledge representation, language understanding, large text corpora



11 Adding imageability features to information displays

 Matthew Chalmers, Robert Ingram, Christoph Pfranger

November 1996 **Proceedings of the 9th annual ACM symposium on User interface software and technology**

Publisher: ACM Press

Full text available: [pdf\(1.42 MB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



Keywords: imageability, information design, navigation, visualization



12 The use of cluster hierarchies in hypertext information retrieval

 D. B. Crouch, C. J. Crouch, G. Andreas

November 1989 **Proceedings of the second annual ACM conference on Hypertext**

Publisher: ACM Press

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index](#)



Full text available: [pdf\(1.05 MB\)](#)

[terms](#)

The graph-traversal approach to hypertext information retrieval is a conceptualization of hypertext in which the structural aspects of the nodes are emphasized. A user navigates through such hypertext systems by evaluating the semantics associated with links between nodes as well as the information contained in nodes. [Fris88] In this paper we describe an hierarchical structure which effectively supports the graphical traversal of a document collection in a hypertext system ...

13 What the query told the link: the integration of hypertext and information retrieval

 Gene Golovchinsky

April 1997 **Proceedings of the eighth ACM conference on Hypertext**

Publisher: ACM Press

Full text available: [pdf\(825.12 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



Keywords: browsing, dynamic hypertext, information exploration, information retrieval, newspaper metaphor, relevance feedback

14 Semantic Clustering of Index Terms

 S. Kumar

October 1968 **Journal of the ACM (JACM)**, Volume 15 Issue 4

Publisher: ACM Press

Full text available: [pdf\(1.30 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



15 Facilitating the creation of a multiple index on graph-described documents by

 transforming their descriptions

Guy William Mineau

December 1993 **Proceedings of the second international conference on Information and knowledge management**

Publisher: ACM Press

Full text available: [pdf\(825.01 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



16 Graphical table of contents

 Xia Lin

April 1996 **Proceedings of the first ACM international conference on Digital libraries**

Publisher: ACM Press

Full text available: [pdf\(978.24 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



17 Semantic and schematic similarities between database objects: a context-based approach

Vipul Kashyap, Amit Sheth

December 1996 **The VLDB Journal — The International Journal on Very Large Data Bases**, Volume 5 Issue 4

Publisher: Springer-Verlag New York, Inc.

Full text available: [pdf\(287.44 KB\)](#) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)



In a multidatabase system, schematic conflicts between two objects are usually of interest

only when the objects have some semantic similarity. We use the concept of *semantic proximity*, which is essentially an *abstraction/mapping* between the domains of the two objects associated with the *context of comparison*. An explicit though partial context representation is proposed and the specificity relationship between contexts is defined. The contexts are organized as a meet semi-l ...

18 Life, death, and lawfulness on the electronic frontier

 James Pitkow, Peter Pirolli

 March 1997 **Proceedings of the SIGCHI conference on Human factors in computing systems**

Publisher: ACM Press

Full text available:  pdf(1.05 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



Keywords: World Wide Web, categorization, clustering, co-citation analysis, hypertext, survival analysis, usage models

19 Queries? Links? Is there a difference?

 Gene Golovchinsky

 March 1997 **Proceedings of the SIGCHI conference on Human factors in computing systems**

Publisher: ACM Press

Full text available:  pdf(1.09 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



Keywords: browsing, dynamic links, hypertext, information exploration, information retrieval, newspaper metaphor

20 Conceptual schema analysis: techniques and applications

 S. Castano, V. De Antonellis, M. G. Fugini, B. Pernici

 September 1998 **ACM Transactions on Database Systems (TODS)**, Volume 23 Issue 3

Publisher: ACM Press

Full text available:  pdf(350.09 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)



The problem of analyzing and classifying conceptual schemas is becoming increasingly important due to the availability of a large number of schemas related to existing applications. The purposes of schema analysis and classification activities can be different: to extract information on intensional properties of legacy systems in order to restructure or migrate to new architectures; to build libraries of reference conceptual components to be used in building new applications in a given domain ...

Keywords: conceptual modeling, reference components, schema classification, schema similarity

Results 1 - 20 of 200

Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)

 **PORTAL**
USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

Search: The ACM Digital Library The Guide

+clustering documents, semantics, similarities

SEARCH

THE ACM DIGITAL LIBRARY

 [Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Found 7,095 of 102,899

Sort results
by[relevance](#) [Save results to a Binder](#)[Try an Advanced Search](#)Display
results[expanded form](#) [Search Tips](#)[Try this search in The ACM Guide](#) [Open results in a new window](#)

Results 21 - 40 of 200

Result page: [previous](#)

1

2

3

4

5

6

7

8

9

10

[next](#)

Best 200 shown

Relevance scale 

- 21** [An interface for navigating clustered document sets returned by queries](#) 

 Robert B. Allen, Pascal Obry, Michael Littman

December 1993 **Proceedings of the conference on Organizational computing systems**

Publisher: ACM Press

Full text available:  [pdf\(571.98 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: clustering, information retrieval, interfaces

- 22** [Enriched knowledge representation for information retrieval](#) 

 F. N. Teskey

November 1987 **Proceedings of the 10th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  [pdf\(630.00 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

In this paper we identify the need for a new theory of information. An information model is developed which distinguishes between data, as directly observable facts, information, as structured collections of data, and knowledge as methods of using information. The model is intended to support a wide range of information systems. In the paper we develop the use of the model for a semantic information retrieval system using the concept of semantic categories. The likely benefits of this area ...

- 23** [Text categorization for multiple users based on semantic features from a machine-](#) 

 [readable dictionary](#)

Elizabeth D. Liddy, Woojin Paik, Edmund S. Yu

July 1994 **ACM Transactions on Information Systems (TOIS)**, Volume 12 Issue 3

Publisher: ACM Press

Full text available:  [pdf\(1.17 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

The text categorization module described here provides a front-end filtering function for the larger DR-LINK text retrieval system [Liddy and Myaeing 1993]. The model evaluates a large incoming stream of documents to determine which documents are sufficiently similar to a profile at the broad subject level to warrant more refined representation and

matching. To accomplish this task, each substantive word in a text is first categorized using a feature set based on the semantic Subject Field ...

Keywords: semantic vectors, subject field coding

24 Document retrieval and text retrieval: An overview of DR-LINK and its approach to document filtering

Elizabeth D. Liddy, Woojin Paik, Edmund S. Yu, Kenneth A. McVearry

March 1993 **Proceedings of the workshop on Human Language Technology HLT '93**

Publisher: Association for Computational Linguistics

Full text available:  pdf(409.85 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

DR-LINK is an information retrieval system, complex in design and processing, with the potential for providing significant advances in retrieval results due to the range and richness of semantic representation done by the various modules in the system. By using a full continuum of linguistic-conceptual processing, DR-LINK has the capability of producing documents which precisely match users' needs. Each of DR-LINK's six processing modules add to the conceptual enhancement of the document and que ...

25 Term clustering of syntactic phrases

 D. D. Lewis, W. B. Croft

December 1989 **Proceedings of the 13th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(1.62 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Term clustering and syntactic phrase formation are methods for transforming natural language text. Both have had only mixed success as strategies for improving the quality of text representations for document retrieval. Since the strengths of these methods are complementary, we have explored combining them to produce superior representations. In this paper we discuss our implementation of a syntactic phrase generator, as well as our preliminary experiments with producing phrase clusters. Th ...

26 Contextualizing the information space in federated digital libraries

 M. P. Papazoglou, J. Hoppenbrouwers

March 1999 **ACM SIGMOD Record**, Volume 28 Issue 1

Publisher: ACM Press

Full text available:  pdf(695.20 KB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

Rapid growth in the volume of documents, their diversity, and terminological variations render federated digital libraries increasingly difficult to manage. Suitable abstraction mechanisms are required to construct meaningful and scalable document clusters, forming a cross-digital library information space for browsing and semantic searching. This paper addresses the above issues, proposes a distributed semantic framework that achieves a logical partitioning of the information space accordi ...

27 Adaptive document clustering

 C. T. Yu, Y. T. Wang, C. H. Chen

June 1985 **Proceedings of the 8th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(465.87 KB) Additional Information: [full citation](#), [references](#), [citations](#)

28 Structured hypertext with domain semantics Weigang Wang, Roy RadaOctober 1998 **ACM Transactions on Information Systems (TOIS)**, Volume 16 Issue 4**Publisher:** ACM PressFull text available:  pdf(593.99 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

One important facet of current hypertext research involves using knowledge-based techniques to develop and maintain document structures. A semantic net is one such technique. However, most semantic-net-based hypertext systems leave the linking consistency of the net to individual users. Users without guidance may accidentally introduce structural and relational inconsistencies in the semantic nets. The relational inconsistency hinders the creation of domain information models. The structura ...

Keywords: graph theory, hypertext models, hypertext structures**29 What if there were desktop access to the computer science literature?** Dennis J. Brueni, Baziley T. Cross, Edward A. Fox, Lenwood S. Heath, Deborah Hix, Lucy T. Nowell, William C. WakeMarch 1993 **Proceedings of the 1993 ACM conference on Computer science****Publisher:** ACM PressFull text available:  pdf(1.18 MB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

What if there were an electronic computer science library? Consider the possibilities of having your favorite publications available within finger's reach. Consider Project Envision, an ongoing effort to build a user-centered database from the computer science literature. This paper describes the project's first year progress, stressing the underlying motivation, user-centered development, and the overall design.

30 A self-organizing semantic map for information retrieval Xia Lin, Dagobert Soergel, Gary MarchioniniSeptember 1991 **Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval****Publisher:** ACM PressFull text available:  pdf(797.66 KB)Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**31 Guessing morphology from terms and corpora** Christian JacqueminJuly 1997 **ACM SIGIR Forum , Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '97**, Volume 31 Issue SI**Publisher:** ACM PressFull text available:  pdf(1.62 MB)Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**32 Distributional clustering of words for text classification** L. Douglas Baker, Andrew Kachites McCallumAugust 1998 **Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval****Publisher:** ACM PressFull text available:  pdf(1.07 MB)Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

33 Information retrieval using a singular value decomposition model of latent semantic structure

 G. W. Furnas, S. Deerwester, S. T. Dumais, T. K. Landauer, R. A. Harshman, L. A. Streeter, K. E. Lochbaum

May 1988 **Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(1.35 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

In a new method for automatic indexing and retrieval, implicit higher-order structure in the association of terms with documents is modeled to improve estimates of term-document association, and therefore the detection of relevant documents on the basis of terms found in queries. Singular-value decomposition is used to decompose a large term by document matrix into 50 to 150 orthogonal factors from which the original matrix can be approximated by linear combination; both documents and terms ...

34 Cat-a-Cone: an interactive interface for specifying searches and viewing retrieval results using a large category hierarchy

 Marti A. Hearst, Chandu Karadi

July 1997 **ACM SIGIR Forum , Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '97**, Volume 31 Issue SI

Publisher: ACM Press

Full text available:  pdf(3.95 MB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

35 Approximating matrix multiplication for pattern recognition tasks

Edith Cohen, David D. Lewis

January 1997 **Proceedings of the eighth annual ACM-SIAM symposium on Discrete algorithms**

Publisher: Society for Industrial and Applied Mathematics

Full text available:  pdf(1.10 MB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

36 InfoCrystal: a visual tool for information retrieval & management

 Anselm Spoerri

December 1993 **Proceedings of the second international conference on Information and knowledge management**

Publisher: ACM Press

Full text available:  pdf(999.84 KB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

37 HypIR: a hypertext-based approach to information retrieval

 Fazli Can, Yuan-Ming Lee

March 1993 **Proceedings of the 1993 ACM/SIGAPP symposium on Applied computing: states of the art and practice**

Publisher: ACM Press

Full text available:  pdf(770.39 KB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: browsing, hypertext, information retrieval, inverted files

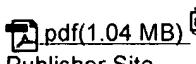
38 Towards the automatic identification of adjectival scales: clustering adjectives according to meaning

Vasileios Hatzivassiloglou, Kathleen R. McKeown

June 1993 **Proceedings of the 31st annual meeting on Association for Computational Linguistics**

Publisher: Association for Computational Linguistics

Full text available:



[pdf\(1.04 MB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)
[Publisher Site](#)

In this paper we present a method to group adjectives according to their meaning, as a first step towards the automatic identification of adjectival scales. We discuss the properties of adjectival scales and of groups of semantically related adjectives and how they imply sources of linguistic knowledge in text corpora. We describe how our system exploits this linguistic knowledge to compute a measure of similarity between two adjectives, using statistical techniques and without having access to ...

39 Interactive clustering for navigating in hypermedia systems

 Sougata Mukherjea, James D. Foley, Scott E. Hudson

September 1994 **Proceedings of the 1994 ACM European conference on Hypermedia technology**

Publisher: ACM Press

Full text available:



[pdf\(918.11 KB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

This paper talks about clustering related nodes of an overview diagram to reduce its complexity and size. This is because although overview diagrams are useful for helping the user to navigate in a hypermedia system, for any real-world system these become too complicated and large to be really useful. Both structure-based and content-based clustering are used. Since the nodes can be related to each other in different ways, depending on the situation different clustered views will be useful. ...

Keywords: clustering, information visualization, navigation, overview diagrams

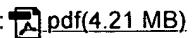
40 Fast detection of communication patterns in distributed executions

Thomas Kunz, Michiel F. H. Seuren

November 1997 **Proceedings of the 1997 conference of the Centre for Advanced Studies on Collaborative research**

Publisher: IBM Press

Full text available:



[pdf\(4.21 MB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Understanding distributed applications is a tedious and difficult task. Visualizations based on process-time diagrams are often used to obtain a better understanding of the execution of the application. The visualization tool we use is Poet, an event tracer developed at the University of Waterloo. However, these diagrams are often very complex and do not provide the user with the desired overview of the application. In our experience, such tools display repeated occurrences of non-trivial commun ...

Results 21 - 40 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads: [!\[\]\(e462f608b89f421f9d905728e26f6429_img.jpg\) Adobe Acrobat](#) [!\[\]\(d23a1e673062c7f9dd7f31b0bf3952ee_img.jpg\) QuickTime](#) [!\[\]\(b6decc5949267b59236e06e16f30320f_img.jpg\) Windows Media Player](#) [!\[\]\(1302b7c44ebac71a8fe8000e9986c651_img.jpg\) Real Player](#)

 **PORTAL**
USPTO

[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

Search: The ACM Digital Library The Guide
 +clustering documents, semantics, similarities

SEARCH

THE ACM DIGITAL LIBRARY

 [Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Found 7,095 of 102,899

Sort results by

relevance

 [Save results to a Binder](#)[Try an Advanced Search](#)

Display results

expanded form

 [Search Tips](#)[Try this search in The ACM Guide](#) [Open results in a new window](#)Results 41 - 60 of 200 Result page: [previous](#) [1](#) [2](#) **3** [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale 

- 41** [From latent semantics to spatial hypertext—an integrated approach](#) 

 Chaomei Chen, Mary Czerwinski

May 1998 **Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space---structure in hypermedia systems: links, objects, time and space---structure in hypermedia systems**

Publisher: ACM Press

Full text available:  [pdf\(1.32 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

- 42** [Special issue on word sense disambiguation: Topical clustering of MRD senses based on information retrieval techniques](#) 

Jen Nan Chen, Jason S. Chang

March 1998 **Computational Linguistics**, Volume 24 Issue 1

Publisher: MIT Press

Full text available:   [pdf\(2.08 MB\)](#) [HTML](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)
[Publisher Site](#)

This paper describes a heuristic approach capable of automatically clustering senses in a machine-readable dictionary (MRD). Including these clusters in the MRD-based lexical database offers several positive benefits for word sense disambiguation (WSD). First, the clusters can be used as a coarser sense division, so unnecessarily fine sense distinction can be avoided. The clustered entries in the MRD can also be used as materials for supervised training to develop a WSD system. Furthermore, if t ...

- 43** [Exploration of text collections with hierarchical feature maps](#) 

 Dieter Merkl

July 1997 **ACM SIGIR Forum , Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '97**, Volume 31 Issue SI

Publisher: ACM Press

Full text available:  [pdf\(1.65 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

- 44** [Towards an information logic](#) 

C. J. van Rijsbergen

May 1989 **ACM SIGIR Forum , Proceedings of the 12th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '89**, Volume 23 Issue SI

Publisher: ACM Press

Full text available:  pdf(820.63 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

'Probability is expectation founded upon partial knowledge.' (Boole, 1854) Information retrieval based on stored program electronic computers has been an active area of research since the time these machines were invented. It is therefore somewhat surprising that even now no formal computational model for IR exists. There is no well-defined logic to describe information retrieval, and there is no proof or model theory to talk about the truths of IR. This paper argues t ...

45 User-oriented document clustering: a framework for learning in information retrieval 

 J. S. Deogun, V. V. Raghavan

September 1986 **Proceedings of the 9th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(776.71 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

In information retrieval, cluster analysis is an important tool employed to enhance both efficiency and effectiveness of the retrieval process. Most clustering algorithms have difficulty in reflecting the closeness of documents as perceived by the user. A two phase scheme for document clustering, whose results reflect the "conceptual" clusters that are perceived by the user of the retrieval system, is proposed. Since the clusters obtained by this scheme are not characterized in ...

46 Hypertext databases and data mining 

 Soumen Chakrabarti

June 1999 **ACM SIGMOD Record , Proceedings of the 1999 ACM SIGMOD international conference on Management of data SIGMOD '99**, Volume 28 Issue 2

Publisher: ACM Press

Full text available:  pdf(151.84 KB) Additional Information: [full citation](#), [abstract](#), [index terms](#)

The volume of unstructured text and hypertext data far exceeds that of structured data. Text and hypertext are used for digital libraries, product catalogs, reviews, newsgroups, medical reports, customer service reports, and the like. Currently measured in billions of dollars, the worldwide internet activity is expected to reach a trillion dollars by 2002. Database researchers have kept some cautious distance from this action. The goal of this tutorial is to expose database researchers to t ...

47 ABSTRACTS (Chosen by G. Salton or V. Raghavan from 1983 issues of journals in the retrieval area) 

 January 1986 **ACM SIGIR Forum**, Volume 19 Issue 1-4

Publisher: ACM Press

Full text available:  pdf(2.06 MB) Additional Information: [full citation](#), [abstract](#)

This article outlines basic videodisc and optical disk technology. Both optical and capacitance videodisc technology are described. Optical disk technology as a mass digital image and data storage device is defined and briefly compared with other established information storage media including magnetic tape and microforms. The article includes a look into the future of videodisc and optical disk.

48 Redundancy: helping semantic disambiguation 

Caroline Barrière

August 1998 **Proceedings of the 17th international conference on Computational**

**linguistics - Volume 1 , Proceedings of the 36th annual meeting on
Association for Computational Linguistics - Volume 1****Publisher:** Association for Computational Linguistics , Association for Computational LinguisticsFull text available:  pdf(556.98 KB) Publisher SiteAdditional Information: [full citation](#), [abstract](#), [references](#)

Redundancy is a good thing, at least in a learning process. To be a good teacher you must say what you are going to say, say it, then say what you have just said. Well, three times is better than one. To acquire and learn knowledge from text for building a lexical knowledge base, we need to find a source of information that states facts, and repeats them a few times using slightly different sentence structures. A technique is needed for gathering information from that source and identify the red ...

49 Inference networks for document retrieval H. Turtle, W. B. CroftDecember 1989 **Proceedings of the 13th annual international ACM SIGIR conference
on Research and development in information retrieval****Publisher:** ACM PressFull text available:  pdf(1.65 MB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The use of inference networks to support document retrieval is introduced. A network-based retrieval model is described and compared to conventional probabilistic and Boolean models.

50 Query processing in a multimedia document system Elisa Bertino, Fausto Rabitti, Simon GibbsJanuary 1988 **ACM Transactions on Information Systems (TOIS)**, Volume 6 Issue 1**Publisher:** ACM PressFull text available:  pdf(2.94 MB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Query processing in a multimedia document system is described. Multimedia documents are information objects containing formatted data, text, image, graphics, and voice. The query language is based on a conceptual document model that allows the users to formulate queries on both document content and structure. The architecture of the system is outlined, with focus on the storage organization in which both optical and magnetic devices can coexist. Query processing and the different strategies ...

51 Discovering similar resources by content part-linking Brad Perry, Wesley W. ChuJanuary 1997 **Proceedings of the sixth international conference on Information and
knowledge management****Publisher:** ACM PressFull text available:  pdf(1.10 MB)Additional Information: [full citation](#), [references](#), [index terms](#)**52 Applications of linear algebra in information retrieval and hypertext analysis** Jon Kleinberg, Andrew TomkinsMay 1999 **Proceedings of the eighteenth ACM SIGMOD-SIGACT-SIGART symposium
on Principles of database systems****Publisher:** ACM PressFull text available:  pdf(1.12 MB)Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

53 Clustering in large graphs and matrices

P. Drineas, Alan Frieze, Ravi Kannan, Santosh Vempala, V. Vinay

January 1999 **Proceedings of the tenth annual ACM-SIAM symposium on Discrete algorithms**

Publisher: Society for Industrial and Applied Mathematics

Full text available:  pdf(927.53 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

**54 A study of probability kinematics in information retrieval**

 F. Crestani, C. J. van Rijsbergen

July 1998 **ACM Transactions on Information Systems (TOIS)**, Volume 16 Issue 3

Publisher: ACM Press

Full text available:  pdf(1.12 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)



We analyze the kinematics of probabilistic term weights at retrieval time for different Information Retrieval models. We present four models based on different notions of probabilistic retrieval. Two of these models are based on classical probability theory and can be considered as prototypes of models long in use in Information Retrieval, like the Vector Space Model and the Probabilistic Model. The two other models are based on a logical technique of evaluating the probability of a conditi ...

Keywords: logical imaging, probabilistic modeling, probabilistic retrieval

55 Polyrepresentation of information needs and semantic entities: elements of a cognitive theory for information retrieval interaction

Peter Ingwersen

August 1994 **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: Springer-Verlag New York, Inc.

Full text available:  pdf(1.16 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#), [review](#)

**56 Methods for information server selection**

 David Hawking, Paul Thistlewaite

January 1999 **ACM Transactions on Information Systems (TOIS)**, Volume 17 Issue 1

Publisher: ACM Press

Full text available:  pdf(283.76 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)



The problem of using a broker to select a subset of available information servers in order to achieve a good trade-off between document retrieval effectiveness and cost is addressed. Server selection methods which are capable of operating in the absence of global information, and where servers have no knowledge of brokers, are investigated. A novel method using Lightweight Probe queries (LWP method) is compared with several methods based on data from past query processing, while Random and ...

Keywords: Lightweight Probe queries, information servers, network servers, server ranking, server selection, text retrieval

57 A performance comparison of object and relational databases using the Sun Benchmark

 Joshua Duhl, Craig Damon
 January 1988 **ACM SIGPLAN Notices , Conference proceedings on Object-oriented programming systems, languages and applications OOPSLA '88**, Volume 23 Issue 11

Publisher: ACM Press

Full text available:  pdf(1.01 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

A general concern about object-oriented systems has been whether or not they are able to meet the performance demands required to be useful for the development of significant production software systems. Attempts to evaluate this assertion have been hampered by a lack of meaningful performance benchmarks that compare database operations across different kinds of databases. In this paper, we utilize the Sun Benchmark [Rube87] as a means for assessing the performance of an object d ...

58 A look back and a look forward

 K. Sparck Jones
 May 1988 **Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(1.57 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

This paper is in two parts, following the suggestion that I first comment on my own past experience in information retrieval, and then present my views on the present and future.

59 A document retrieval system for man-machine interaction

 G. Salton
 January 1964 **Proceedings of the 1964 19th ACM national conference**

Publisher: ACM Press

Full text available:  pdf(1.07 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

An automatic document retrieval system, programmed for the IBM 7094, is described. The system is designed to process English texts and search requests, and uses statistical, syntactic and semantic procedures for the analysis of information and the identification of relevant items. The operations are planned around a central supervisor, which in turn calls on the various subroutines, as desired. This organization makes it possible to alter both the processing sequences and the mat ...

60 Role of verbs in document analysis

Judith Klavans, Min-Yen Kan
 August 1998 **Proceedings of the 17th international conference on Computational linguistics - Volume 1 , Proceedings of the 36th annual meeting on Association for Computational Linguistics - Volume 1**

Publisher: Association for Computational Linguistics , Association for Computational Linguistics

Full text available:  pdf(706.80 KB)

 Publisher Site

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

We present results of two methods for assessing the event profile of news articles as a function of verb type. The unique contribution of this research is the focus on the role of verbs, rather than nouns. Two algorithms are presented and evaluated, one of which is shown to accurately discriminate documents by type and semantic properties, i.e. the event profile. The initial method, using WordNet (Miller et al. 1990), produced multiple cross-classification of articles, primarily due to the bushy ...

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.
[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)

 [Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

Search: The ACM Digital Library The Guide

+clustering documents, semantics, similarities

THE ACM DIGITAL LIBRARY

 [Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

Found 7,095 of 102,899

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Sort results by

[relevance](#) [Save results to a Binder](#)[Try an Advanced Search](#)

Display results

[expanded form](#) [Search Tips](#)[Try this search in The ACM Guide](#) [Open results in a new window](#)

Results 61 - 80 of 200

Result page: [previous](#)

1

2

3

4

5

6

7

8

9

10

next

Best 200 shown

Relevance scale **61 Papers: An automatic clustering of articles using dictionary definitions**

Fumiyo Fukumoto, Yoshimi Suzuki

August 1996 **Proceedings of the 16th conference on Computational linguistics - Volume 1****Publisher:** Association for Computational LinguisticsFull text available:  [pdf\(542.17 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)

In this paper, we propose a statistical approach for clustering of articles using on-line dictionary definitions. One of the characteristics of our approach is that every sense of word in articles is automatically disambiguated using dictionary definitions. The other is that in order to cope with the problem of a phrasal lexicon, linking which links words with their semantically similar words in articles is introduced in our method. The results of experiments demonstrate the effectiveness of the ...

62 Project summaries: Robust text processing and information retrieval

Tomek Strzalkowski

March 1994 **Proceedings of the workshop on Human Language Technology HLT '94****Publisher:** Association for Computational LinguisticsFull text available:  [pdf\(102.12 KB\)](#) Additional Information: [full citation](#), [abstract](#)

The general objective of this research has been the enhancement of traditional key-word based statistical methods of document retrieval with advanced natural language processing techniques. In the work to date the focus has been on obtaining a better representation of document contents by extracting representative phrases from syntactically preprocessed text and devising suitable weighting schemes for different types of terms. In addition, statistical clustering methods have been developed that ...

63 Automatic noun classification by using Japanese-English word pairs

Naomi Inoue

June 1991 **Proceedings of the 29th annual meeting on Association for Computational Linguistics****Publisher:** Association for Computational LinguisticsFull text available:  [pdf\(479.46 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#) [Publisher Site](#)

This paper describes a method of classifying semantically similar nouns. The approach is based on the "distributional hypothesis". Our approach is characterized by distinguishing

among senses of the same word in order to resolve the "polysemy" issue. The classification result demonstrates that our approach is successful.

64 Generation and search of clustered files

 G. Salton, A. Wong

December 1978 **ACM Transactions on Database Systems (TODS)**, Volume 3 Issue 4

Publisher: ACM Press

Full text available:  pdf(1.78 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

A classified, or clustered file is one where related, or similar records are grouped into classes, or clusters of items in such a way that all items within a cluster are jointly retrievable. Clustered files are easily adapted to broad and narrow search strategies, and simple file updating methods are available. An inexpensive file clustering method applicable to large files is given together with appropriate file search methods. An abstract model is then introduced to predict the retrieval ...

Keywords: automatic classification, cluster searching, clustered files, fast classification, file organization, probabilistic models

65 Computational Methods for Intelligent Information Access

 Michael W. Berry, Susan T. Dumais, Todd A. Letsche

December 1995 **Proceedings of the 1995 ACM/IEEE conference on Supercomputing (CDROM) - Volume 00 Supercomputing '95**

Publisher: ACM Press, IEEE Computer Society

Full text available:  pdf(375.60 KB)

 html(3.13 KB)

Additional Information: [full citation](#), [references](#), [citations](#)

66 Abstraction mechanisms in CLU

 Barbara Liskov, Alan Snyder, Russell Atkinson, Craig Schaffert

August 1977 **Communications of the ACM**, Volume 20 Issue 8

Publisher: ACM Press

Full text available:  pdf(1.43 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

CLU is a new programming language designed to support the use of abstractions in program construction. Work in programming methodology has led to the realization that three kinds of abstractions—procedural, control, and especially data abstractions—are useful in the programming process. Of these, only the procedural abstraction is supported well by conventional languages, through the procedure or subroutine. CLU provides, in addition to procedures, novel linguistic mechanisms th ...

Keywords: control abstractions, data abstractions, data types, programming languages, programming methodology, separate compilation

67 Output ranking methodology for document-clustering-based Boolean retrieval systems

 Tadeusz Radecki

June 1985 **Proceedings of the 8th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(649.81 KB)

Additional Information: [full citation](#), [references](#), [citations](#)

68 Multi-media document representation and retrieval Esen Ozkarahan, Fazli CanApril 1991 **Proceedings of the 19th annual conference on Computer Science****Publisher:** ACM PressFull text available: .pdf(1.10 MB)Additional Information: [full citation](#), [references](#), [citations](#)**69 Special issue on word sense disambiguation: Introduction to the special issue on word sense disambiguation: the state of the art**

Nancy Ide, Jean Véronis

March 1998 **Computational Linguistics**, Volume 24 Issue 1**Publisher:** MIT PressFull text available: .pdf(3.44 MB) Additional Information: [full citation](#), [references](#), [citations](#)[Publisher Site](#)**70 Automating the assignment of submitted manuscripts to reviewers** Susan T. Dumais, Jakob NielsenJune 1992 **Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval****Publisher:** ACM PressFull text available: .pdf(1.39 MB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The 117 manuscripts submitted for the Hypertext '91 conference were assigned to members of the review committee, using a variety of automated methods based on information retrieval principles and Latent Semantic Indexing. Fifteen reviewers provided exhaustive ratings for the submitted abstracts, indicating how well each abstract matched their interests. The automated methods do a fairly good job of assigning relevant papers for review, but they are still somewhat poorer than ...

71 Evaluation of an inference network-based retrieval model Howard Turtle, W. Bruce CroftJuly 1991 **ACM Transactions on Information Systems (TOIS)**, Volume 9 Issue 3**Publisher:** ACM PressFull text available: .pdf(2.40 MB)Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#), [review](#)**Keywords:** document retrieval, inference networks, network retrieval models**72 A model of a document-clustering-based information retrieval system with a Boolean search request formulation**

Tadeusz Radecki

June 1980 **Proceedings of the 3rd annual ACM conference on Research and development in information retrieval****Publisher:** Butterworth & Co.Full text available: .pdf(562.73 KB)Additional Information: [full citation](#), [references](#), [citations](#)

73 Information filtering: the computation of similarities in large corpora of legal texts Erich Schweighofer, Werner Winiwarter, Dieter Merkl May 1995 **Proceedings of the 5th international conference on Artificial intelligence and law****Publisher:** ACM PressFull text available:  pdf(748.19 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**74 The exploration of legal text corpora with hierarchical neural networks: a guided tour in public international law** Dieter Merkl, Erich SchweighoferJune 1997 **Proceedings of the 6th international conference on Artificial intelligence and law****Publisher:** ACM PressFull text available:  pdf(1.12 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**75 Exploiting clustering and phrases for context-based information retrieval** Peter G. Anick, Shivakumar Vaithyanathan July 1997 **ACM SIGIR Forum , Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '97**, Volume 31 Issue SI**Publisher:** ACM PressFull text available:  pdf(1.55 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**76 Retrieving documents by plausible inference: a preliminary study** W. B. Croft, T. J. Lucia, P. R. Cohen May 1988 **Proceedings of the 11th annual international ACM SIGIR conference on Research and development in information retrieval****Publisher:** ACM PressFull text available:  pdf(1.13 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Choosing an appropriate document representation and search strategy for document retrieval has been largely guided by achieving good average performance instead of optimizing the results for each individual query. A model of retrieval based on plausible inference gives us a different perspective and suggests that techniques should be found for combining multiple sources of evidence (or search strategies) into an overall assessment of a document's relevance, rather than attempting to pick a ...

77 Visualizing document authorship using n-grams and latent semantic indexing Ian M. Soboroff, Charles K. Nicholas, James M. Kukla, David S. Ebert November 1997 **Proceedings of the 1997 workshop on New paradigms in information visualization and manipulation****Publisher:** ACM PressFull text available:  pdf(495.64 KB)  ps(2.35 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**Keywords:** clustering, indexing methods, information filtering, linguistic processing

- 78 [The newspaper image database: empirical supported analysis of users' typology and word association clusters](#) 

Susanne Ornager

July 1995 **Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(787.36 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#) 

- 79 [An evaluation of phrasal and clustered representations on a text categorization task](#) 

David D. Lewis

June 1992 **Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(1.22 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Syntactic phrase indexing and term clustering have been widely explored as text representation techniques for text retrieval. In this paper we study the properties of phrasal and clustered indexing languages on a text categorization task, enabling us to study their properties in isolation from query interpretation issues. We show that optimal effectiveness occurs when using only a small proportion of the indexing terms available, and that effectiveness peaks at a higher feature set size and ...

- 80 [Advantages of query biased summaries in information retrieval](#) 

Anastasios Tombros, Mark Sanderson

August 1998 **Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(1.12 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Results 61 - 80 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) **4** [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)



[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

Search: The ACM Digital Library The Guide

+clustering documents, semantics, similarities

SEARCH

THE ACM DIGITAL LIBRARY

[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

Found 7,095 of 102,899

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Sort results
by

[relevance](#)

[Save results to a Binder](#)

[Try an Advanced Search](#)

Display
results

[expanded form](#)

[Search Tips](#)

[Try this search in The ACM Guide](#)

[Open results in a new window](#)

Results 81 - 100 of 200

Result page: [previous](#)

[1](#)

[2](#)

[3](#)

[4](#)

[5](#)

[6](#)

[7](#)

[8](#)

[9](#)

[10](#)

[next](#)

Best 200 shown

Relevance scale

81 [VAGUE: a user interface to relational databases that permits vague queries](#)

Amihai Motro

July 1988 **ACM Transactions on Information Systems (TOIS)**, Volume 6 Issue 3

Publisher: ACM Press

Full text available: [pdf\(2.16 MB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

A specific query establishes a rigid qualification and is concerned only with data that match it precisely. A vague query establishes a target qualification and is concerned also with data that are close to this target. Most conventional database systems cannot handle vague queries directly, forcing their users to retry specific queries repeatedly with minor modifications until they match data that are satisfactory. This article describes a system called VAGUE that can handle vague queries ...

82 [Organization of clustered files for consecutive retrieval](#)

J S. Deogun, V V. Raghavan, T K.W. Tsou

December 1984 **ACM Transactions on Database Systems (TODS)**, Volume 9 Issue 4

Publisher: ACM Press

Full text available: [pdf\(1.79 MB\)](#)

Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#), [review](#)

This paper studies the problem of storing single-level and multilevel clustered files. Necessary and sufficient conditions for a single-level clustered file to have the consecutive retrieval property (CRP) are developed. A linear time algorithm to test the CRP for a given clustered file and to identify the proper arrangement of objects, if CRP exists, is presented. For the single-level clustered files that do not have CRP, it is shown that the problem of identifying a storage organization w ...

83 [Analytical versus empirical evaluation of spatial displays](#)

Mountaz Hascoët

April 1998 **CHI '98 conference summary on Human factors in computing systems**

Publisher: ACM Press

Full text available: [pdf\(269.28 KB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: analytic versus empirical evaluation, browsing, spatial layout algorithms,

visualization

84 Cluster-based scalable network services

 Armando Fox, Steven D. Gribble, Yatin Chawathe, Eric A. Brewer, Paul Gauthier
October 1997 **ACM SIGOPS Operating Systems Review , Proceedings of the sixteenth ACM symposium on Operating systems principles SOSP '97**, Volume 31 Issue 5

Publisher: ACM Press

Full text available:  pdf(2.42 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

85 Automatic subject indexing using an associative neural network

 Yi-Ming Chung, William M. Pottenger, Bruce R. Schatz
May 1998 **Proceedings of the third ACM conference on Digital libraries**

Publisher: ACM Press

Full text available:  pdf(1.26 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

86 ER model clustering as an aid for user communication and documentation in database design

 Toby J. Teorey, Guangping Wei, Deborah L. Bolton, John A. Koenig
August 1989 **Communications of the ACM**, Volume 32 Issue 8

Publisher: ACM Press

Full text available:  pdf(1.14 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Entity-relationship clustering promotes the simplicity that is vital for fast end-user comprehension, as well as the complexity at a more detailed level to satisfy the database designer's need for extended semantic expression in the conceptual model.

87 Selected IR-Related Dissertation Abstracts

 September 1991 **ACM SIGIR Forum**, Volume 25 Issue 2

Publisher: ACM Press

Full text available:  pdf(2.75 MB) Additional Information: [full citation](#), [abstract](#)

The following are citations selected by title and abstract as being related to Information Retrieval (IR), resulting from a computer search, using BRS Information Technologies, of the Dissertation Abstracts Online database produced by University Microfilms International (UMI). Included are UMI order number, title, author, degree, year, institution; number of pages, one or more Dissertation Abstracts International (DAI) subject descriptors chosen by the author, and abstract. Unless otherwise spec ...

88 Dynamic document processing

 Gerard Salton
July 1972 **Communications of the ACM**, Volume 15 Issue 7

Publisher: ACM Press

Full text available:  pdf(1.07 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

The current role of computers in automatic document processing is briefly outlined, and some reasons are given why the early promise of library automation and of the mechanization of documentation processes has not been fulfilled. A new dynamic document environment is then outlined in which clustered files are searched and information is retrieved following an interactive user-controlled search process. Methods

are described for an automatic query modification based on user needs ...

Keywords: automatic indexing, automatic search and retrieval, collection growth, collection retirement, feedback search, iterative searching, mechanized library processing

89 Document detection: DR-LINK system: phase I summary

Elizabeth D. Liddy, Sung H. Myaeng

September 1993 **Proceedings of a workshop on held at Fredericksburg, Virginia: September 19-23, 1993**

Publisher: Association for Computational Linguistics

Full text available: [pdf\(1.66 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)

The underlying principle of the DR-LINK System is that retrieval must be at the conceptual level, not the word level. That is, a successful retrieval system must retrieve on the basis of what people **mean** in their query, not just what they say in their query. The same is true of documents - their representation needs to capture the content at the conceptual level of expression. To accomplish this human-like goal, DR-LINK aims to represent and match documents and queries at all of the avail ...

90 Incorporating different search models into one document retrieval system

W. Bruce Croft

May 1981 **ACM SIGIR Forum , Proceedings of the 4th annual international ACM SIGIR conference on Information storage and retrieval: theoretical issues in information retrieval SIGIR '81**, Volume 16 Issue 1

Publisher: ACM Press

Full text available: [pdf\(580.78 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Many effective search strategies derived from different models are available for document retrieval systems. However, it does not appear that there is a single most effective strategy. Instead, different strategies perform optimally under different conditions. This paper outlines the design of an adaptive document retrieval system that chooses the best search strategy for a particular situation and user. In order to be able to support a variety of search strategies, a general network representat ...

91 Special issue on knowledge representation

Ronald J. Brachman, Brian C. Smith
February 1980 **ACM SIGART Bulletin**, Issue 70

Publisher: ACM Press

Full text available: [pdf\(13.13 MB\)](#) Additional Information: [full citation](#), [abstract](#)

In the fall of 1978 we decided to produce a special issue of the SIGART Newsletter devoted to a survey of current knowledge representation research. We felt that there were twe useful functions such an issue could serve. First, we hoped to elicit a clear picture of how people working in this subdiscipline understand knowledge representation research, to illuminate the issues on which current research is focused, and to catalogue what approaches and techniques are currently being developed. Secon ...

92 Optimal determination of user-oriented clusters

J. Deogun, V. Raghavan
November 1987 **Proceedings of the 10th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available: [pdf\(632.90 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

User-oriented clustering schemes enable the classification of documents based upon the

user perception of the similarity between documents, rather than on some similarity function presumed by the designer to represent the user criteria. In this paper, an enhancement of such a clustering scheme is presented. This is accomplished by the formulation of the user-oriented clustering as a function-optimization problem. The problem formulated is termed the Boundary Selection Problem (BSP). Heurist ...

93 Incremental clustering and dynamic information retrieval

 Moses Charikar, Chandra Chekuri, Tomás Feder, Rajeev Motwani

May 1997 **Proceedings of the twenty-ninth annual ACM symposium on Theory of computing**

Publisher: ACM Press

Full text available:  pdf(1.58 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



94 Automatic retrieval and clustering of similar words

Dekang Lin

August 1998 **Proceedings of the 17th international conference on Computational linguistics - Volume 2 , Proceedings of the 36th annual meeting on Association for Computational Linguistics - Volume 2**

Publisher: Association for Computational Linguistics , Association for Computational Linguistics

Full text available:  pdf(554.99 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

 Publisher Site



Bootstrapping semantics from text is one of the greatest challenges in natural language learning. We first define a word similarity measure based on the distributional pattern of words. The similarity measure allows us to construct a thesaurus using a parsed corpus. We then present a new evaluation methodology for the automatically constructed thesaurus. The evaluation results show that the thesaurus is significantly closer to WordNet than Roget Thesaurus is.

95 Recent Studies in Automatic Text Analysis and Document Retrieval

 G. Salton

April 1973 **Journal of the ACM (JACM)**, Volume 20 Issue 2

Publisher: ACM Press

Full text available:  pdf(1.56 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)



Many experts in mechanized text processing now agree that useful automatic language analysis procedures are largely unavailable and that the existing linguistic methodologies generally produce disappointing results. An attempt is made in the present study to identify those automatic procedures which appear most effective as a replacement for the missing language analysis. A series of computer experiments is described, designed to simulate a conventional document retrieval environ ...

96 Dimensions of meaning

H. Schütze

December 1992 **Proceedings of the 1992 ACM/IEEE conference on Supercomputing**

Publisher: IEEE Computer Society Press

Full text available:  pdf(964.54 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



97 A dynamic cluster maintenance system for information retrieval

 F. Can, E. Ozkarahan

November 1987 **Proceedings of the 10th annual international ACM SIGIR conference**



on Research and development in information retrieval

Publisher: ACM Press

Full text available: [!\[\]\(f15da8627380db409bac161a6cb03047_img.jpg\) pdf\(937.52 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Partitioning by clustering of very large databases is a necessity to reduce the space/time complexity of retrieval operations. However, the contemporary and modern retrieval environments demand dynamic maintenance of clusters. A new cluster maintenance strategy is proposed and its similarity/stability characteristics, cost analysis, and retrieval behavior in comparison with unclustered and completely reclustered database environments have been examined by means of a series of experi ...

98 Information storage and retrieval: a survey and functional description



Jack Minker
September 1977 **ACM SIGIR Forum**, Volume 12 Issue 2

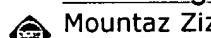
Publisher: ACM Press

Full text available: [!\[\]\(71f9d84356bec8aef5a389e4c1b7f2e3_img.jpg\) pdf\(5.14 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)

Information Storage and Retrieval (IS&R) encompasses a broad scope of topics ranging from basic techniques for accessing data to sophisticated approaches for the analysis of natural language text and the deduction of information. Within the field, three general areas of investigation can be distinguished not only by their subject matter but also by the types of individuals presently interested in them:(1) Document retrieval,(2) Generalized data management, and(3) Question-answering.A functional ...

Keywords: automatic indexing, data management, data structures, deductive search, information retrieval, natural language, problem solving, question-answering, relational data systems, theorem proving

99 Accessing hyperdocuments through interactive dynamic maps



Mountaz Zizi, Michel Beaudouin-Lafon
September 1994 **Proceedings of the 1994 ACM European conference on Hypermedia technology**

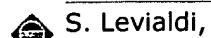
Publisher: ACM Press

Full text available: [!\[\]\(24f6c2ce6f1490d1d5ae4f3dfdf89611_img.jpg\) pdf\(1.39 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

We propose a new navigation paradigm based on a spatial metaphor to help users access and navigate within large sets of documents. This metaphor is implemented by a computer artifact called an Interactive Dynamic Map (IDM). An IDM plays a role similar to the role of a real map with respect to physical space. Two types of IDMs are computed from the documents: Topic IDMs represent the semantic contents of a set of documents while Document IDMs visualize a subset of documents ...

Keywords: information retrieval, interaction paradigm, maps, navigation, visualization

100 The interface of the future



S. Levialdi, A. N. Badre, M. Chalmers, P. Copeland, P. Mussio, C. Solomon
June 1994 **Proceedings of the workshop on Advanced visual interfaces**

Publisher: ACM Press

Full text available: [!\[\]\(8fa7794540bcfe454084a5cf2d2c66a1_img.jpg\) pdf\(663.89 KB\)](#) Additional Information: [full citation](#), [index terms](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.
[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)



[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

Search: The ACM Digital Library The Guide

+clustering documents, semantics, similarities

SEARCH

THE ACM DIGITAL LIBRARY

[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

Found 7,095 of 102,899

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Sort results by

[relevance](#)

[Save results to a Binder](#)

Try an [Advanced Search](#)

Display results

[expanded form](#)

[Search Tips](#)

Try this search in [The ACM Guide](#)

[Open results in a new window](#)

Results 101 - 120 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) **6** [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale

101 Personalized information structures II: hyperstructure hotlists

Scott R. Tilley, Walter M. Lamia

February 1996 **Proceedings of the 13th annual international conference on Systems documentation: emerging from chaos: solutions for the growing complexity of our jobs**

Publisher: ACM Press

Full text available: [pdf\(945.18 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



102 Automated hoarding for mobile computers

Geoffrey H. Kuenning, Gerald J. Popek

October 1997 **ACM SIGOPS Operating Systems Review , Proceedings of the sixteenth ACM symposium on Operating systems principles SOSP '97**, Volume 31 Issue 5

Publisher: ACM Press

Full text available: [pdf\(2.05 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



103 HieNet: a user-centered approach for automatic link generation

Daniel T. Chang

December 1993 **Proceedings of the fifth ACM conference on Hypertext**

Publisher: ACM Press

Full text available: [pdf\(1.14 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



Keywords: SGML, link apprentice, link generation, links, vector space model

104 Statistical language modeling: Augmenting lexicons automatically: clustering semantically related adjectives

Kathleen McKeown, Vasileios Hatzivassiloglou

March 1993 **Proceedings of the workshop on Human Language Technology HLT '93**

Publisher: Association for Computational Linguistics

Full text available: [pdf\(549.98 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)



Our work focuses on identifying various types of lexical data in large corpora through statistical analysis. In this paper, we present a method for grouping adjectives according to their meaning, as a step towards the automatic identification of adjectival scales. We describe how our system exploits two sources of linguistic knowledge in a corpus to compute a measure of similarity between two adjectives, using statistical techniques and a clustering algorithm for grouping. We evaluate the signif ...

105 Protofoil: storing and finding the information worker's paper documents in an electronic file cabinet

 Ramana Rao, Stuart K. Card, Walter Johnson, Leigh Klotz, Randall H. Trigg
April 1994 **Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence**

Publisher: ACM Press

Full text available:  pdf(1.38 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



Keywords: ad hoc information work, document imaging, filing of paper documents, information retrieval, paper user interface

106 Description of EDCS technology clusters

 September 1997 **ACM SIGSOFT Software Engineering Notes**, Volume 22 Issue 5

Publisher: ACM Press

Full text available:  pdf(1.14 MB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)



Evolutionary Systems are those that are capable of accomodating change over an extended system lifetime with reduced risk and cost/schedule impact. Most of our complex defense systems depend on software for their successful operation and, as a result, the software in those systems is the primary vehicle for adapting to change. The EDCS (Evolutionary Design of Complex Software) Program is providing for the development and experimental application of new software technologies which can enable signi ...

107 An Analysis of Some Graph Theoretical Cluster Techniques

 J. Gary Augustson, Jack Minker
October 1970 **Journal of the ACM (JACM)**, Volume 17 Issue 4

Publisher: ACM Press

Full text available:  pdf(1.29 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)



Several graph theoretic cluster techniques aimed at the automatic generation of thesauri for information retrieval systems are explored. Experimental cluster analysis is performed on a sample corpus of 2267 documents. A term-term similarity matrix is constructed for the 3950 unique terms used to index the documents. Various threshold values, T, are applied to the similarity matrix to provide a series of binary threshold matrices. The corresponding graph of each binary thres ...

108 Office documents on a database kernel—filing, retrieval, and archiving

 P. Zabback, H. B. Paul, U. Deppisch
March 1990 **ACM SIGOIS Bulletin , Proceedings of the ACM SIGOIS and IEEE CS TC-OA conference on Office information systems**, Volume 11 Issue 2-3

Publisher: ACM Press

Full text available:  pdf(1.24 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)



One of the main component of integrated office systems is the large central filing system.

It efficiently stores, retrieves and searches office documents containing text, images, graphics, data and voice. We propose to implement a filing system on top of the Darmstadt database system (DASDBS), which is designed as a data management kernel for both standard and non-standard applications. This paper investigates the choice of appropriate storage structures for the filing system objects and th ...

109 Toward a logical/physical theory of spreadsheet modeling

 Tomás Isakowitz, Shimon Schocken, Henry C. Lucas

January 1995 **ACM Transactions on Information Systems (TOIS)**, Volume 13 Issue 1

Publisher: ACM Press

Full text available:  pdf(2.76 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

In spite of the increasing sophistication and power of commercial spreadsheet packages, we still lack a formal theory or a methodology to support the construction and maintenance of spreadsheet models. Using a dual logical/physical perspective, we identify four principal components that characterize any spread sheet model: schema, data, editorial, and binding. We present a factoring algorithm for identifying and extracting these components ...

Keywords: model management

110 Indexing hypertext documents in context

 Guy A. Boy

September 1991. **Proceedings of the third annual ACM conference on Hypertext**

Publisher: ACM Press

Full text available:  pdf(901.26 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

111 Using clustering and visualization for refining the results of a WWW image search engine



Sougata Mukherjea, Kyoji Hirata, Yoshinori Hara

November 1998 **Proceedings of the 1998 workshop on New paradigms in information visualization and manipulation**

Publisher: ACM Press

Full text available:  pdf(1.61 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: World Wide Web, image search engine, query result visualization

112 Abstracting of legal cases: the SALOMON experience

 Marie-Francine Moens, Caroline Uyttendaele, Jos Dumortier

June 1997 **Proceedings of the 6th international conference on Artificial intelligence and law**

Publisher: ACM Press

Full text available:  pdf(1.24 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

113 FastMap: a fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets

 Christos Faloutsos, King-IP Lin

May 1995 **ACM SIGMOD Record , Proceedings of the 1995 ACM SIGMOD international conference on Management of data SIGMOD '95**, Volume 24 Issue 2

Publisher: ACM Press

Full text available:  pdf(1.21 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

A very promising idea for fast searching in traditional and multimedia databases is to map objects into points in k -d space, using k feature-extraction functions, provided by a domain expert [25]. Thus, we can subsequently use highly fine-tuned spatial access methods (SAMs), to answer several types of queries, including the 'Query By Example' type (which translates to a range query); the 'all pairs' query (which translates to a spatial join [8]); the nearest-neighbor or best-match ...

114 Scatter/gather browsing communicates the topic structure of a very large text

 **collection**

Peter Pirolli, Patricia Schank, Marti Hearst, Christine Diehl

April 1996 **Proceedings of the SIGCHI conference on Human factors in computing systems: common ground**

Publisher: ACM Press

Full text available:  pdf(1.23 MB) 

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#) [html\(47.35 KB\)](#)

Keywords: Scatter/Gather, browsing, clustering, information retrieval

115 Improving text retrieval for the routing problem using latent semantic indexing

David Hull

August 1994 **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: Springer-Verlag New York, Inc.

Full text available:  pdf(883.83 KB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#), [review](#)

116 Adaptive information retrieval: using a connectionist representation to retrieve and

 **learn about documents**

R. K. Belew

May 1989 **ACM SIGIR Forum , Proceedings of the 12th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '89**, Volume 23 Issue SI

Publisher: ACM Press

Full text available:  pdf(1.19 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

AIR represents a connectionist approach to the task of information retrieval. The system uses relevance feedback from its users to change its representation of authors, index terms and documents so that, over time, AIR improves at its task. The result is a representation of the consensual meaning of keywords and documents shared by some group of users. The central focus goal of this paper is to use our experience with AIR to highlight those characteristics of connectionist ...

117 Automatic linking of thesauri

 S. Amba, N. Narasimhamurthi, Kevin C. O'Kane, Philip M. Turner

August 1996 **Proceedings of the 19th annual international ACM SIGIR conference on**

Research and development in information retrieval**Publisher:** ACM PressFull text available:  pdf(613.76 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**118 Document classification using a finite mixture model** 

Hang Li, Kenji Yamanishi

July 1997 Proceedings of the 35th annual meeting on Association for Computational Linguistics , Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics**Publisher:** Association for Computational Linguistics , Association for Computational LinguisticsFull text available:  pdf(769.56 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#) [Publisher Site](#)

We propose a new method of classifying documents into categories. We define for each category a *finite mixture model* based on *soft clustering* of words. We treat the problem of classifying documents as that of conducting statistical hypothesis testing over finite mixture models, and employ the EM algorithm to efficiently estimate parameters in a finite mixture model. Experimental results indicate that our method outperforms existing methods.

119 Selected Information Retrieval Abstracts  Gerry SaltonNovember 1990 **ACM SIGIR Forum**, Volume 24 Issue 3**Publisher:** ACM PressFull text available:  pdf(844.03 KB) Additional Information: [full citation](#)**120 Structuring and visualising the WWW by generalised similarity analysis**  Chaomei ChenApril 1997 **Proceedings of the eighth ACM conference on Hypertext****Publisher:** ACM PressFull text available:  pdf(962.97 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: WWW, information visualisation, pathfinder networks, sequential behavioural patterns, structural analysis

Results 101 - 120 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) **6** [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)



[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

Search: The ACM Digital Library The Guide

+clustering documents, semantics, similarities

SEARCH

THE ACM DIGITAL LIBRARY

 [Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

Found 7,095 of 102,899

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Sort results
by

relevance

 [Save results to a Binder](#)

Display
results

expanded form

 [Search Tips](#)
 Open results in a new window

Try an [Advanced Search](#)
Try this search in [The ACM Guide](#)

Results 121 - 140 of 200 Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) **7** [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale 

121 [Final report of the ANSI/X3/SPARC DBS-SG relational database task group](#)

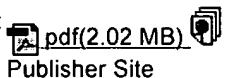
 July 1982 **ACM SIGMOD Record**, Volume 12 Issue 4

Publisher: ACM Press

Full text available:  [pdf\(4.69 MB\)](#)

Publisher: MIT Press

Full text available:



Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

[Publisher Site](#)

This paper addresses the problem of how to identify the intended meaning of individual words in unrestricted texts, without necessarily having access to complete representations of sentences. To discriminate senses, an understander can consider a diversity of information, including syntactic tags, word frequencies, collocations, semantic context, role-related expectations, and syntactic restrictions. However, current approaches make use of only small subsets of this information. Here we will des ...

- 125 [Automatic generation of "hyper-paths" in information retrieval systems: a stochastic and an incremental algorithms](#)

Alain Lelu

September 1991 **Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available: [pdf\(981.55 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



- 126 [Research in information extraction & document detection: A context vector-based self organizing map for information visualization](#)

David A. Rushall, Marc R. Ilgen

May 1996 **Proceedings of a workshop on held at Vienna, Virginia: May 6-8, 1996**

Publisher: Association for Computational Linguistics

Full text available: [pdf\(708.10 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)



HNC Software, Inc. has developed a system called DOCUVERSE for visualizing the information content of large textual corpora. The system is built around two separate neural network methodologies: context vectors and self organizing maps. Context vectors (CVs) are high dimensional information representations that encode the semantic content of the textual entities they represent. Self organizing maps (SOMs) are capable of transforming an input, high dimensional signal space into a much lower (usu ...

- 127 [Atomic data abstractions in a distributed collaborative editing system](#)

Irene Greif, Robert Seliger, William E. Weihl

January 1986 **Proceedings of the 13th ACM SIGACT-SIGPLAN symposium on Principles of programming languages**

Publisher: ACM Press

Full text available: [pdf\(2.52 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)



This paper describes our experience implementing CES, a distributed Collaborative Editing System written in Argus, a language that includes facilities for managing long-lived distributed data. Argus provides *<i>atomic actions</i>* which simplify the handling of concurrency and failures, and mechanisms for implementing *<i>atomic data types</i>* which ensure serializability and recoverability of actions that use them. This paper focuses on the support for atomicity in Argus ...

- 128 [Experiments on the determination of the relationships between terms](#)

Vijay V. Raghavan, C. T. Yu

June 1979 **ACM Transactions on Database Systems (TODS)**, Volume 4 Issue 2

Publisher: ACM Press

Full text available: [pdf\(1.52 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)



The retrieval effectiveness of an automatic method that uses relevance judgments for the

determination of positive as well as negative relationships between terms is evaluated. The term relationships are incorporated into the retrieval process by using a generalized similarity function that has a term match component, a positive term relationship component, and a negative term relationship component. Two strategies, query partitioning and query clustering, for the evaluation of the effectiv ...

Keywords: antonym, document retrieval, feedback, pseudoclassification, semantics, statistical discrimination, synonym, term associations, thesaurus

129 A study of retrospective and on-line event detection

 Yiming Yang, Tom Pierce, Jaime Carbonell

August 1998 **Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(1.05 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



130 Selected IR-Related Dissertation Abstracts

 February 1992 **ACM SIGIR Forum**, Volume 26 Issue 1

Publisher: ACM Press

Full text available:  pdf(2.24 MB) Additional Information: [full citation](#)



131 Selected IR-Related Dissertation Abstracts

 March 1993 **ACM SIGIR Forum**, Volume 27 Issue 1

Publisher: ACM Press

Full text available:  pdf(2.24 MB) Additional Information: [full citation](#), [abstract](#)



The following are citations selected by title and abstract as being related to Information Retrieval (IR), resulting from a computer search, using BRS Information Technologies, of the Dissertation Abstracts Online database produced by University Microfilms International (UMI). Included are UMI order number, title, author, degree, year, institution; number of pages, and abstract. Unless otherwise specified, paper or microform copies of dissertations may be ordered from University Microfilms Inter ...

132 Customizing information capture and access

 Daniela Rus, Devika Subramanian

January 1997 **ACM Transactions on Information Systems (TOIS)**, Volume 15 Issue 1

Publisher: ACM Press

Full text available:  pdf(1.26 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)



This article presents a customizable architecture for software agents that capture and access information in large, heterogeneous, distributed electronic repositories. The key idea is to exploit underlying structure at various levels of granularity to build high-level indices with task-specific interpretations. Information agents construct such indices and are configured as a network of reusable modules called structure detectors and segmenters. We illustrate our architectu ...

Keywords: information gathering, software agents, table recognition

133

The personal electronic program guide—towards the pre-selection of individual TV



 programs

Michael Ehrmantraut, Theo Härdter, Hartmut Wittig, Ralf Steinmetz
November 1996 **Proceedings of the fifth international conference on Information and knowledge management**

Publisher: ACM Press

Full text available:  pdf(923.87 KB) Additional Information: [full citation](#), [references](#), [index terms](#)

134 Concept based query expansion 

 Yonggang Qiu, Hans-Peter Frei
July 1993 **Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(1.05 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Query expansion methods have been studied for a long time - with debatable success in many instances. In this paper we present a probabilistic query expansion model based on a similarity thesaurus which was constructed automatically. A similarity thesaurus reflects domain knowledge about the particular collection from which it is constructed. We address the two important issues with query expansion: the selection and the weighting of additional search terms. In contrast to earlier methods, ...

135 Special issue: AI in engineering 

 D. Sriram, R. Joobhani
April 1985 **ACM SIGART Bulletin**, Issue 92

Publisher: ACM Press

Full text available:  pdf(8.79 MB) Additional Information: [full citation](#), [abstract](#)

The papers in this special issue were compiled from responses to the announcement in the July 1984 issue of the SIGART newsletter and notices posted over the ARPAnet. The interest being shown in this area is reflected in the sixty papers received from over six countries. About half the papers were received over the computer network.

136 Composite objects revisited 

 Won Kim, Elisa Bertino, Jorge F. Garza
June 1989 **ACM SIGMOD Record , Proceedings of the 1989 ACM SIGMOD international conference on Management of data SIGMOD '89**, Volume 18 Issue 2

Publisher: ACM Press

Full text available:  pdf(1.57 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

In object-oriented systems, an object may recursively reference any number of other objects. The references, however, do not capture any special relationships between objects. An important semantic relationship which may be superimposed on a reference is the IS-PART-OF relationship between a pair of objects. A set of objects related by the IS-PART-OF relationship is collectively called a composite object. An earlier paper [KIM87b] presented a model of composite objects which has ...

137 File clustering using naming conventions for legacy systems 

Nicolas Anquetil, Timothy Lethbridge
November 1997 **Proceedings of the 1997 conference of the Centre for Advanced Studies on Collaborative research**

Publisher: IBM Press

Full text available:  pdf(243.06 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Decomposing complex software systems into conceptually independent subsystems represents a significant software engineering activity that receives considerable research attention. Most of the research in this domain deals with the source code; trying to cluster together files which are conceptually related. In this paper we propose using a more informal source of information: file names. We present an experiment which shows that file naming convention is the best file clustering criteria for the ...

138 Probability kinematics in information retrieval F. Crestani, C. J. van Rijsbergen July 1995 **Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval****Publisher:** ACM PressFull text available:  [pdf\(919.37 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**139 Lexical ambiguity and information retrieval** Robert Krovetz, W. Bruce Croft April 1992 **ACM Transactions on Information Systems (TOIS)**, Volume 10 Issue 2**Publisher:** ACM PressFull text available:  [pdf\(2.00 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Lexical ambiguity is a pervasive problem in natural language processing. However, little quantitative information is available about the extent of the problem or about the impact that it has on information retrieval systems. We report on an analysis of lexical ambiguity in information retrieval test collections and on experiments to determine the utility of word meanings for separating relevant from nonrelevant documents. The experiments show that there is considerable ambiguity even in a s ...

Keywords: disambiguation, document retrieval, semantically based search, word senses

140 Conference abstracts January 1977 **Proceedings of the 5th annual ACM computer science conference****Publisher:** ACM PressFull text available:  [pdf\(3.14 MB\)](#) Additional Information: [full citation](#), [abstract](#), [index terms](#)

One problem in computer program testing arises when errors are found and corrected after a portion of the tests have run properly. How can it be shown that a fix to one area of the code does not adversely affect the execution of another area? What is needed is a quantitative method for assuring that new program modifications do not introduce new errors into the code. This model considers the retest philosophy that every program instruction that could possibly be reached and tested from the ...

Results 121 - 140 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)


[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)
Search: The ACM Digital Library The Guide

+clustering documents, semantics, similarities

THE ACM DIGITAL LIBRARY
[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

Found 7,095 of 102,899

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Sort results by

[relevance](#)
 [Save results to a Binder](#)

Display results

[expanded form](#)
 [Search Tips](#)
 [Open results in a new window](#)
[Try an Advanced Search](#)
[Try this search in The ACM Guide](#)

Results 141 - 160 of 200

Result page: [previous](#)

1

2

3

4

5

6

7

8

9

10

[next](#)

Best 200 shown

Relevance scale

141 Status report of the graphic standards planning committee

Computer Graphics staff

August 1979 **ACM SIGGRAPH Computer Graphics**, Volume 13 Issue 3

Publisher: ACM Press

Full text available: [pdf\(15.01 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#)
142 Abstracts of Articles in the Information Retrieval Area Selected by Gerard Salton

 September 1986 **ACM SIGIR Forum**, Volume 21 Issue 1-2

Publisher: ACM Press

Full text available: [pdf\(1.10 MB\)](#) Additional Information: [full citation](#)
143 Information retrieval theory and design based on a model of the user's concept relations

Matthew B. Koll

June 1980 **Proceedings of the 3rd annual ACM conference on Research and development in information retrieval**

Publisher: Butterworth & Co.

Full text available: [pdf\(753.08 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#)
144 Squibs and discussions: Integrating general-purpose and corpus-based verb classification

Roberto Basili, Paola Velardi, Maria Teresa Pazienza

December 1996 **Computational Linguistics**, Volume 22 Issue 4

Publisher: MIT Press

Full text available: [pdf\(611.52 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#)
[Publisher Site](#)

145 A comparison of classifiers and document representations for the routing problem


 Hinrich Schütze, David A. Hull, Jan O. Pedersen
July 1995 **Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: ACM Press
Full text available:  pdf(1.16 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

146 Distributed systems - programming and management: On remote procedure call 

Patrícia Gomes Soares
November 1992 **Proceedings of the 1992 conference of the Centre for Advanced Studies on Collaborative research - Volume 2**
Publisher: IBM Press
Full text available:  pdf(4.52 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

The Remote Procedure Call (RPC) paradigm is reviewed. The concept is described, along with the backbone structure of the mechanisms that support it. An overview of works in supporting these mechanisms is discussed. Extensions to the paradigm that have been proposed to enlarge its suitability, are studied. The main contributions of this paper are a standard view and classification of RPC mechanisms according to different perspectives, and a snapshot of the paradigm in use today and of goals for t ...

147 Experiments in automatic statistical thesaurus construction 

 Carolyn J. Crouch, Bokyung Yang
June 1992 **Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: ACM Press
Full text available:  pdf(741.19 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

A well constructed thesaurus has long been recognized as a valuable tool in the effective operation of an information retrieval system. This paper reports the results of experiments designed to determine the validity of an approach to the automatic construction of global thesauri (described originally by Crouch in [1] and [2] based on a clustering of the document collection. The authors validate the approach by showing that the use of thesauri generated by this method results in substantial ...

148 Hypertext paradigm in the field of information retrieval: a neural approach 

 Alain Lelu, Claire Francois
December 1993 **Proceedings of the ACM conference on Hypertext**
Publisher: ACM Press
Full text available:  pdf(894.79 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: cluster analysis, graphic user interface, hypertext, information retrieval, neural networks

149 Constructing information systems based on schema reuse 

 Wen-Syan Li, Richard D. Holowczak
November 1996 **Proceedings of the fifth international conference on Information and knowledge management**
Publisher: ACM Press
Full text available:  pdf(945.90 KB) Additional Information: [full citation](#), [references](#), [index terms](#)

150 [Integrated document caching and prefetching in storage hierarchies based on Markov-chain predictions](#)

Achim Kraiss, Gerhard Weikum

August 1998 **The VLDB Journal — The International Journal on Very Large Data Bases**, Volume 7 Issue 3

Publisher: Springer-Verlag New York, Inc.

Full text available:  pdf(603.01 KB) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

Large multimedia document archives may hold a major fraction of their data in tertiary storage libraries for cost reasons. This paper develops an integrated approach to the vertical data migration between the tertiary, secondary, and primary storage in that it reconciles speculative prefetching, to mask the high latency of the tertiary storage, with the replacement policy of the document caches at the secondary and primary storage level, and also considers the interaction of these policies with ...

Keywords: Caching, Markov chains, Performance, Prefetching, Scheduling, Stochastic modeling, Tertiary storage

151 [Concept features in Re:Agent, an intelligent Email agent](#)

 Gary Boone

May 1998 **Proceedings of the second international conference on Autonomous agents**

Publisher: ACM Press

Full text available:  pdf(1.07 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

152 [Evaluating document retrieval in patent database: a preliminary report](#)

 Mark Osborn, Tomek Strzalkowski, Mihnea Marinescu

January 1997 **Proceedings of the sixth international conference on Information and knowledge management**

Publisher: ACM Press

Full text available:  pdf(811.55 KB) Additional Information: [full citation](#), [references](#), [index terms](#)

153 [COOL: system support for distributed programming](#)

 Rodger Lea, Christian Jacquemot, Eric Pillevesse

September 1993 **Communications of the ACM**, Volume 36 Issue 9

Publisher: ACM Press

Full text available:  pdf(3.45 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#), [review](#)

Keywords: concurrency, concurrent object-oriented programming

154 [Query expansion using local and global document analysis](#)

 Jinxi Xu, W. Bruce Croft

August 1996 **Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(722.09 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

155 Automatic selection of class labels from a thesaurus for an effective semantic tagging of corpora

Alessandro Cucchiarelli, Paola Velardi

March 1997 **Proceedings of the fifth conference on Applied natural language processing**

Publisher: Morgan Kaufmann Publishers Inc.

Full text available:  pdf(708.88 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

 Publisher Site

It is widely accepted that tagging text with semantic information would improve the quality of lexical learning in corpus-based NLP methods. However available on-line taxonomies are rather entangled and introduce an unnecessary level of ambiguity. The noise produced by the redundant number of tags often overrides the advantage of semantic tagging. In this paper we propose an automatic method to select from WordNet a subset of domain-appropriate categories that effectively reduce the overambiguity ...

156 Semantic database modeling: survey, applications, and research issues

 Richard Hull, Roger King

September 1987 **ACM Computing Surveys (CSUR)**, Volume 19 Issue 3

Publisher: ACM Press

Full text available:  pdf(5.42 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Most common database management systems represent information in a simple record-based format. Semantic modeling provides richer data structuring capabilities for database applications. In particular, research in this area has articulated a number of constructs that provide mechanisms for representing structurally complex interrelations among data typically arising in commercial applications. In general terms, semantic modeling complements work on knowledge representation (in artificial int ...

157 Flatland: new dimensions in office whiteboards

 Elizabeth D. Mynatt, Takeo Igarashi, W. Keith Edwards, Anthony LaMarca

May 1999 **Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit**

Publisher: ACM Press

Full text available:  pdf(1.11 MB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Flatland is an augmented whiteboard interface designed for informal office work. Our research investigates approaches to building an augmented whiteboard in the context of continuous, long term office use. In particular, we pursued three avenues of research based on input from user studies: techniques for the management of space on the board, the ability to flexibly apply behaviors to support varied application semantics, and mechanisms for managing history on the board. Unlike some p ...

Keywords: Flatland, light-weight interaction, pen-based computing, ubiquitous computing, whiteboards

158 A semidiscrete matrix decomposition for latent semantic indexing information retrieval

 Tamara G. Kolda, Dianne P. O'Leary

October 1998 **ACM Transactions on Information Systems (TOIS)**, Volume 16 Issue 4

Publisher: ACM Press

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index](#)

Full text available: [pdf\(201.61 KB\)](#)

[terms](#), [review](#)

The vast amount of textual information available today is useless unless it can be effectively and efficiently searched. The goal in information retrieval is to find documents that are relevant to a given user query. We can represent a document collection by a matrix whose (i, j) entry is nonzero only if the i th term appears in the j th document; thus each document corresponds to a column vector. The query is also represented as a column $v \dots$

Keywords: data mining, latent semantic indexing, semidiscrete decomposition, singular-value decomposition, text retrieval

159 Multi-level transaction management for complex objects: implementation,

[performance](#), [parallelism](#)

Gerhard Weikum, Christof Hasse

October 1993 **The VLDB Journal — The International Journal on Very Large Data Bases**, Volume 2 Issue 4

Publisher: Springer-Verlag New York, Inc.

Full text available: [pdf\(2.83 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Multi-level transactions are a variant of open-nested transactions in which the subtransactions correspond to operations at different levels of a layered system architecture. They allow the exploitation of semantics of high-level operations to increase concurrency. As a consequence, undoing a transaction requires compensation of completed subtransactions. In addition, multi-level recovery methods must take into consideration that high-level operations are not necessarily atomic if multiple pages ...

Keywords: atomicity, complex objects, inter- and intratransaction parallelism, multi-level transactions, performance, persistence, recovery

160 Local Feedback in Full-Text Retrieval Systems

 R. Attar, A. S. Fraenkel

July 1977 **Journal of the ACM (JACM)**, Volume 24 Issue 3

Publisher: ACM Press

Full text available: [pdf\(1.36 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Results 141 - 160 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) **8** [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)


[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)

 Search: The ACM Digital Library The Guide

+clustering documents, semantics, similarities

THE ACM DIGITAL LIBRARY
[Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

 Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Found 7,095 of 102,899

 Sort results
by

relevance

 [Save results to a Binder](#)
[Try an Advanced Search](#)

 Display
results

expanded form

 [Search Tips](#)
[Try this search in The ACM Guide](#)
 [Open results in a new window](#)

 Results 161 - 180 of 200 Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale

161 An object-oriented data model for distributed office applications

E. Bertino, M. Negri, G. Pelagatti, L. Sbattella

 March 1990 **ACM SIGOIS Bulletin , Proceedings of the ACM SIGOIS and IEEE CS TC-OA conference on Office information systems**, Volume 11 Issue 2-3

Publisher: ACM Press

 Full text available: [pdf\(1.19 MB\)](#)

 Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The object-oriented paradigm is becoming very popular for database applications and several object-oriented DBMSs have been developed. A basic notion in this paradigm is the inheritance hierarchy that allows the users to define objects and the associated operations starting from already defined objects. However, in database applications the inheritance hierarchy must provide a conceptual modeling function, in addition to the re-usability function. Another important requirement is to provide ...

162 A survey of structured and object-oriented software specification methods and
[techniques](#)

Roel Wieringa

 December 1998 **ACM Computing Surveys (CSUR)**, Volume 30 Issue 4

Publisher: ACM Press

 Full text available: [pdf\(605.26 KB\)](#)

 Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

This article surveys techniques used in structured and object-oriented software specification methods. The techniques are classified as techniques for the specification of external interaction and internal decomposition. The external specification techniques are further subdivided into techniques for the specification of functions, behavior, and communication. After surveying the techniques, we summarize the way they are used in structured and object-oriented methods and indicate ways in w...

Keywords: languages

163 Semiautomatic generation of glossary links: a practical solution

Hermann Kaindl, Stefan Kramer, Papa Samba Niang Diallo

 February 1999 **Proceedings of the tenth ACM Conference on Hypertext and hypermedia : returning to our diverse roots: returning to our diverse**

roots**Publisher:** ACM PressFull text available:  pdf(2.14 MB)Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**Keywords:** WWW, authoring, automatic link generation, glossary links, hypertext

- 164 Knowledge acquisition from texts: using an automatic clustering method based on noun-modifier relationship** 

Houssem AssadiJuly 1997 **Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics , Proceedings of the 35th annual meeting on Association for Computational Linguistics****Publisher:** Association for Computational Linguistics , Association for Computational LinguisticsFull text available:  pdf(290.54 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#) Publisher Site

We describe the early stage of our methodology of knowledge acquisition from technical texts. First, a partial morpho-syntactic analysis is performed to extract "candidate terms". Then, the knowledge engineer, assisted by an automatic clustering tool, builds the "conceptual fields" of the domain. We focus on this conceptual analysis stage, describe the data prepared from the results of the morpho-syntactic analysis and show the results of the clustering module and their interpretation. We found ...

- 165 Launching the new era** 

 Kazuhiro Fuchi, Robert Kowalski, Koichi Furukawa, Kazunori Ueda, Ken Kahn, Takashi Chikayama, Evan Tick

March 1993 **Communications of the ACM**, Volume 36 Issue 3**Publisher:** ACM PressFull text available:  pdf(3.45 MB)Additional Information: [full citation](#), [references](#), [index terms](#), [review](#)

- 166 A multilevel approach to intelligent information filtering: model, system, and evaluation** 

 J. Mostafa, S. Mukhopadhyay, M. Palakal, W. Lam

October 1997 **ACM Transactions on Information Systems (TOIS)**, Volume 15 Issue 4**Publisher:** ACM PressFull text available:  pdf(610.01 KB)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

In information-filtering environments, uncertainties associated with changing interests of the user and the dynamic document stream must be handled efficiently. In this article, a filtering model is proposed that decomposes the overall task into subsystem functionalities and highlights the need for multiple adaptation techniques to cope with uncertainties. A filtering system, SIFTER, has been implemented based on the model, using established techniques in information retrieval and artificial ...

Keywords: automated document representation, information filtering, user modeling

- 167 Natural language processing for information retrieval** 

 David D. Lewis, Karen Spärck Jones

January 1996 **Communications of the ACM**, Volume 39 Issue 1

Publisher: ACM Press

Full text available:  pdf(602.45 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



168 Dissertation Abstracts

 May 1988 **ACM SIGIR Forum**, Volume 22 Issue 3-4

Publisher: ACM Press

Full text available:  pdf(1.85 MB) Additional Information: [full citation](#)



169 Technical reports

 SIGACT News Staff

January 1980 **ACM SIGACT News**, Volume 12 Issue 1

Publisher: ACM Press

Full text available:  pdf(5.28 MB) Additional Information: [full citation](#)



170 Experiments with a component theory of probabilistic information retrieval based on

 single terms as document components

K. L. Kwok

October 1990 **ACM Transactions on Information Systems (TOIS)**, Volume 8 Issue 4

Publisher: ACM Press

Full text available:  pdf(1.84 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

A component theory of information retrieval using single content terms as component for queries and documents was reviewed and experimented with. The theory has the advantages of being able to (1) bootstrap itself, that is, define initial term weights naturally based on the fact that items are self relevant; (2) make use of within-item term frequencies; (3) account for query-focused and document-focused indexing and retrieval strategies cooperatively; and (4) allow for component-specific fe ...

Keywords: document-focused and query-focused relevance feedback, indexing and retrieval, inverse collection term frequency weighting, inverse document frequency weighting, probabilistic indexing, probabilistic retrieval, ranking and weighting of composite objects



171 Special issue on persistent object systems: Orthogonally persistent object systems

Malcolm Atkinson, Ronald Morrison

July 1995 **The VLDB Journal — The International Journal on Very Large Data Bases**, Volume 4 Issue 3

Publisher: Springer-Verlag New York, Inc.

Full text available:  pdf(5.02 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Persistent Application Systems (PASs) are of increasing social and economic importance. They have the potential to be long-lived, concurrently accessed, and consist of large bodies of data and programs. Typical examples of PASs are CAD/CAM systems, office automation, CASE tools, software engineering environments, and patient-care support systems in hospitals. Orthogonally persistent object systems are intended to provide improved support for the design, construction, maintenance, and operation o ...

Keywords: database programming languages, orthogonal persistence, persistent application systems, persistent programming languages

172 Special issue on natural language generation: Generating natural language summaries from multiple on-line sources

Dragomir R. Radev, Kathleen R. McKeown
September 1998 **Computational Linguistics**, Volume 24 Issue 3

Publisher: MIT Press

Full text available:



[pdf\(2.36 MB\)](#)



Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

[Publisher Site](#)

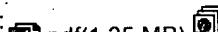
We present a methodology for summarization of news about current events in the form of briefings that include appropriate background (historical) information. The system that we developed, SUMMONS, uses the output of systems developed for the DARPA Message Understanding Conferences to generate summaries of multiple documents on the same or related events, presenting similarities and differences, contradictions, and generalizations among sources of information. We describe the various components ...

173 Special issue on word sense disambiguation: Similarity-based word sense disambiguation

Yael Karov, Shimon Edelman
March 1998 **Computational Linguistics**, Volume 24 Issue 1

Publisher: MIT Press

Full text available:



[pdf\(1.25 MB\)](#)



Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

[Publisher Site](#)

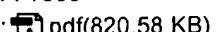
We describe a method for automatic word sense disambiguation using a text corpus and a machine-readable dictionary (MRD). The method is based on word similarity and context similarity measures. Words are considered similar if they appear in similar contexts; contexts are similar if they contain similar words. The circularity of this definition is resolved by an iterative, converging process, in which the system learns from the corpus' a set of typical usages for each of the senses of the polysemous ...

174 Fast and quasi-natural language search for gigabytes of Chinese texts

◆ Lee-Feng Chien
July 1995 **Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:



[pdf\(820.58 KB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

175 Partitioned multiagent systems in information oriented domains

◆ Claudia V. Goldman, Jeffrey S. Rosenschein
April 1999 **Proceedings of the third annual conference on Autonomous Agents**

Publisher: ACM Press

Full text available:



[pdf\(1.05 MB\)](#)

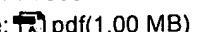
Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

176 Method combination for document filtering

◆ David A. Hull, Jan O. Pedersen, Hinrich Schütze
August 1996 **Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:



[pdf\(1.00 MB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

177 Answering English questions by computer: a survey R. F. SimmonsJanuary 1965 **Communications of the ACM**, Volume 8 Issue 1**Publisher:** ACM PressFull text available:  pdf(2.79 MB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)**178 Abstracts** May 1986 **ACM SIGIR Forum**, Volume 20 Issue 1-4**Publisher:** ACM PressFull text available:  pdf(1.20 MB) Additional Information: [full citation](#), [abstract](#)

The first part of this paper reports a comparative study of the document classifications produced by the use of the single linkage, complete linkage, group average, and Ward clustering methods. Studies of cluster membership and of the effectiveness of cluster searches support previous findings that suggest that the single linkage classifications are rather different from those produced by the other three methods. These latter methods all produce large numbers of small clusters containing just pa ...

179 Database mining challenges for digital libraries Robert GrossmanDecember 1996 **ACM Computing Surveys (CSUR)****Publisher:** ACM PressFull text available:  html(10.82 KB) Additional Information: [full citation](#)**180 Building reusable components in the public administration domain** S. Castano, V. De Antonellis, B. PerniciAugust 1995 **ACM SIGSOFT Software Engineering Notes , Proceedings of the 1995****Symposium on Software reusability SSR '95**, Volume 20 Issue SI**Publisher:** ACM PressFull text available:  pdf(727.94 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The paper proposes methods and tools for building reusable components from families of Information System conceptual schemas, based on the identification of similar components in different schemas, and on their engineering into normalized descriptions. Clustering and abstraction techniques to help identifying similar components, and techniques to build corresponding reusable components are described in the paper.

Results 161 - 180 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.
[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)

 **PORTAL**
USPTO

Subscribe (Full Service) Register (Limited Service, Free) Login
Search: The ACM Digital Library The Guide
 +clustering documents, semantics, similarities

THE ACM DIGITAL LIBRARY

 [Feedback](#) [Report a problem](#) [Satisfaction survey](#)

Published before June 1999

Found 7,095 of 102,899

Terms used [clustering](#) [documents](#) [semantics](#) [similarities](#)

Sort results by [relevance](#)
 Display results [expanded form](#)

[Save results to a Binder](#)
 [Search Tips](#)
 [Open results in a new window](#)

[Try an Advanced Search](#)
[Try this search in The ACM Guide](#)

Results 181 - 200 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) **10**

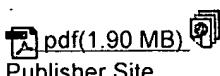
Best 200 shown

Relevance scale **181 Special issue on using large corpora: II: Lexical semantic techniques for corpus analysis**

James Pustejovsky, Peter Anick, Sabine Bergler
 June 1993 **Computational Linguistics**, Volume 19 Issue 2

Publisher: MIT Press

Full text available:

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)[Publisher Site](#)

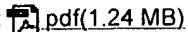
In this paper we outline a research program for computational linguistics, making extensive use of text corpora. We demonstrate how a semantic framework for lexical knowledge can suggest richer relationships among words in text beyond that of simple co-occurrence. The work suggests how linguistic phenomena such as metonymy and polysemy might be exploitable for semantic tagging of lexical items. Unlike with purely statistical collocational analyses, the framework of a semantic theory allows the a ...

182 From documenting design to design by documenting

 Alberto Faro, Daniela Giordano
 October 1997 **Proceedings of the 15th annual international conference on Computer documentation**

Publisher: ACM Press

Full text available:

Additional Information: [full citation](#), [references](#), [index terms](#)

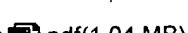
Keywords: case-based reasoning, design memories, scenario-based design

183 Automatic abstracting and indexing—survey and recommendations

 H. P. Edmundson, R. E. Wyllys
 May 1961 **Communications of the ACM**, Volume 4 Issue 5

Publisher: ACM Press

Full text available:

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

In preparation for the widespread use of automatic scanners which will read documents and transmit their contents to other machines for analysis, this report presents a new concept in automatic analysis: the relative-frequency approach to measuring the

significance of words, word groups, and sentences. The relative-frequency approach is discussed in detail, as is its application to problems of automatic indexing and automatic abstracting. Included in the report is a summary of automatic ana ...

184 "Is this document relevant?...probably": a survey of probabilistic models in information retrieval

 Fabio Crestani, Mounia Lalmas, Cornelis J. Van Rijsbergen, Iain Campbell
December 1998 **ACM Computing Surveys (CSUR)**, Volume 30 Issue 4

Publisher: ACM Press

Full text available:  pdf(265.85 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

This article surveys probabilistic approaches to modeling information retrieval. The basic concepts of probabilistic approaches to information retrieval are outlined and the principles and assumptions upon which the approaches are based are presented. The various models proposed in the development of IR are described, classified, and compared using a common formalism. New approaches that constitute the basis of future research are described.

Keywords: information retrieval, probabilistic indexing, probabilistic modeling, probabilistic retrieval, uncertain inference modeling

185 CiteSeer: an autonomous Web agent for automatic retrieval and identification of interesting publications

 Kurt D. Bollacker, Steve Lawrence, C. Lee Giles
May 1998 **Proceedings of the second international conference on Autonomous agents**

Publisher: ACM Press

Full text available:  pdf(1.07 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

Research papers available on the World Wide Web (WWW or Web) are often poorly organized, often exist in forms opaque to search engines (e.g. Postscript), and increase in quantity daily. Significant amounts of time and effort are typically needed in order to find interesting and relevant publications on the Web. We have developed a Web based information agent that assists the user in the process of performing a scientific literature search. Given a set of keywords, the agent uses Web s ...

186 Enhancing design reusability by clustering specifications

 Youwen Ouyang, Doris L. Carver
February 1996 **Proceedings of the 1996 ACM symposium on Applied Computing**

Publisher: ACM Press

Full text available:  pdf(695.67 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: cluster analysis, reusability, software design

187 Enforcing strong object typing in flexible hypermedia

 Pedro Furtado, H. Madeira
May 1998 **Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space---structure in hypermedia systems: links, objects, time and space---structure in hypermedia systems**

Publisher: ACM Press

Full text available: [pdf\(1.16 MB\)](#)

Additional Information: [full citation](#), [references](#), [index terms](#)

188 Visual relevance analysis

 Nikos Pediotakis, Mountaz Hascoët-Zizi

April 1996 **Proceedings of the first ACM international conference on Digital libraries**

Publisher: ACM Press

Full text available: [pdf\(1.61 MB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

189 LyberWorld—a visualization user interface supporting fulltext retrieval

Matthias Hemmje, Clemens Kunkel, Alexander Willett

August 1994 **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: Springer-Verlag New York, Inc.

Full text available: [pdf\(1.31 MB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

190 Personalized information structures

 Scott R. Tilley, Michael J. Whitney, Hausi A. Müller, Margaret-Anne D. Storey

November 1993 **Proceedings of the 11th annual international conference on Systems documentation**

Publisher: ACM Press

Full text available: [pdf\(1.25 MB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: hypertext, information structure, layered graphs, online documentation

191 Information retrieval algorithms: a survey

Prabhakar Raghavan

January 1997 **Proceedings of the eighth annual ACM-SIAM symposium on Discrete algorithms**

Publisher: Society for Industrial and Applied Mathematics

Full text available: [pdf\(908.76 KB\)](#)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

192 Practicing AI with the portable AI lab

 Fabio Baj, Paolo Cattaneo, Mike Rosner

July 1993 **ACM SIGPLAN Lisp Pointers , Proceedings of the third international conference on Lisp users and vendors LUV '93, Volume VI Issue 3**

Publisher: ACM Press

Full text available: [pdf\(697.83 KB\)](#)

Additional Information: [full citation](#), [abstract](#), [index terms](#)

The Portable AI Lab (PAIL) is an integrated collection of modules implementing established AI tools and techniques intended for use as a resource for teaching or learning artificial intelligence (AI). to well illustrated AI concepts by providing a set of incrementally complex demonstration programs, interactive tools to develop new examples, on-line context sensitive documentation with the description of the techniques involved, bibliographic references, and a set of exercises and projects. PAIL ...

193 An application of lexical semantics to knowledge acquisition from corpora

Peter Anick, James Pustejovsky

August 1990 **Proceedings of the 13th conference on Computational linguistics - Volume 2**

Publisher: Association for Computational Linguistics

Full text available: [pdf\(593.19 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

In this paper, we describe a program of research designed to explore how a lexical semantic theory may be exploited for extracting information from corpora suitable for use in Information Retrieval applications. Unlike with purely statistical collocational analyses, the framework of a semantic theory allows the automatic construction of predictions about semantic relationships among words appearing in *collocational systems*. We illustrate the approach for the acquisition of lexical informa ...

Keywords: information retrieval, knowledge acquisition, lexical semantics

194 On the construction of effective vocabularies for information retrieval

G. Salton, Clement T. Yu

November 1973 **ACM SIGIR Forum , ACM SIGPLAN Notices , Proceedings of the 1973 meeting on Programming languages and information retrieval SIGPLAN '73**, Volume 9 , 10 Issue 3 , 1

Publisher: ACM Press

Full text available: [pdf\(837.92 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Natural language query formulations exhibit advantages over artificial language statements since they permit the user to approach the retrieval environment without prior training and without using intermediaries. To obtain adequate retrieval output, it is however necessary to emphasize the good terms and to deemphasize the bad ones. The usefulness of the terms in a natural language vocabulary is first characterized in terms of their frequency distribution over the documents of a collection. The ...

195 Information retrieval using robust natural language processing

Tomek Strzalkowski, Barbara Vauthey

June 1992 **Proceedings of the 30th annual meeting on Association for Computational Linguistics**

Publisher: Association for Computational Linguistics

Full text available: [pdf\(772.67 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

[Publisher Site](#)

We developed a prototype information retrieval system which uses advanced natural language processing techniques to enhance the effectiveness of traditional key-word based document retrieval. The backbone of our system is a statistical retrieval engine which performs automated indexing of documents, then search and ranking in response to user queries. This core architecture is augmented with advanced natural language processing tools which are both robust and efficient. In early experiments, the...

196 A structural view of the Cedar programming environment

Daniel C. Swinehart, Polle T. Zellweger, Richard J. Beach, Robert B. Hagmann

August 1986 **ACM Transactions on Programming Languages and Systems (TOPLAS)**, Volume 8 Issue 4

Publisher: ACM Press

Full text available: [pdf\(6.32 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

This paper presents an overview of the Cedar programming environment, focusing on its overall structure—that is, the major components of Cedar and the way they are organized. Cedar supports the development of programs written in a single programming language, also called Cedar. Its primary purpose is to increase the productivity of

programmers whose activities include experimental programming and the development of prototype software systems for a high-performance personal computer. T ...

197 Knowledge acquisition and representation techniques in scholarly communication

 Brian R. Gaines, Mildred L. G. Shaw

February 1996 **Proceedings of the 13th annual international conference on Systems documentation: emerging from chaos: solutions for the growing complexity of our jobs**

Publisher: ACM Press

Full text available:  pdf(1.13 MB)

Additional Information: [full citation](#), [references](#), [index terms](#)



198 Compiling nested data-parallel programs for shared-memory multiprocessors

 Siddhartha Chatterjee

July 1993 **ACM Transactions on Programming Languages and Systems (TOPLAS)**,

Volume 15 Issue 3

Publisher: ACM Press

Full text available:  pdf(4.17 MB)

Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#), [review](#)



Keywords: compilers, data parallelism, shared-memory multiprocessors

199 Informatics: thesaurus and concordance: Generation of thesaurus in different languages a computer based system

F. J. Devadason

September 1980 **Proceedings of the 8th conference on Computational linguistics**

Publisher: Association for Computational Linguistics

Full text available:  pdf(1.08 MB)

Additional Information: [full citation](#), [abstract](#), [references](#)



The development of the theory of library classification and of subject indexing, for the organisation, storage and retrieval of subjects embodied in documents has a striking parallelism to the search for 'universal forms' and 'deep structure' in language and linguistic studies. The significant contributions of the theories of classification and subject indexing are the subject analysis techniques of Ranganathan and Bhattacharyya's POPSI. A computer based system, for generating an information retr ...

200 Bead: explorations in information visualization

 Matthew Chalmers, Paul Chitton

June 1992 **Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval**

Publisher: ACM Press

Full text available:  pdf(822.89 KB)

Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)



We describe work on the visualization of bibliographic data and, to aid in this task, the application of numerical techniques for multidimensional scaling. Many areas of scientific research involve complex multivariate data. One example of this is Information Retrieval. Document comparisons may be done using a large number of variables. Such conditions do not favour the more well-known methods of visualization and graphical analysis, as it is rarely feasible to map each variable ...

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2006 ACM, Inc.
[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)